Input characteristics and impact on the acquisition of phonological variables in French

Loïc Liégeois, Inès Saddour, Damien Chabanal, Thierry Chanier

Input characteristics and impact on the acquisition of phonological variables in French

Loïc Liégeois[1], Inès Saddour[2], Damien Chabanal[1] et Thierry Chanier[1]
[1]*Clermont Université - LRL*
[2]*Université de Toulouse Le Mirail - Octogone*

Reference:

Liégeois, L., Saddour, I., Chabanal, D. and Chanier, T. Input characteristics and impact on the acquisition of phonological variables in French. Child Language Seminar, 24th-25th June 2013, Manchester.

Introduction

This study is closely related to the ALIPE project. This project has three main objectives. First, we want to collect and structure a corpus of spontaneous interactions between three children and their parents in French as first language. Second, based on the data analyses, we aimed to describe the child directed speech (CDS) and more specifically to give a description of the acquisition of the two phonological variables of liaison and schwa elision. Last, we intended to observe the construction of the variation patterns and propose a scenario for schwa acquisition.

This study consists of three parts. First, we will present a brief literature review concerning CDS and the acquisition of phonological variation. Then, we will explain the methodology used to collect and structure our corpus before presenting our first results concerning the acquisition of schwa elision. First of all, we define the two phonological variables studied in the ALIPE project.

Liaison and schwa elision: definitions

The liaison is the phenomenon of producing a consonant that is even graphically present and normally mute at the end of a word (called Word1) preceding a second one (called Word2) beginning with a vowel. When the Word1 is produced alone or before a word beginning with a consonant, this consonant is mute. For example:

- "*trois ours*" = [tr waz uʀs](three bears) but "*trois*" [tr wa] and "*ours*" [uʀs].

We can distinguish two types of liaison:

- categorical liaisons, that are already produced by any speakers in any situations,
- and variable liaisons whose production depends on several criteria such as the context of interaction and the geographical origin of the speaker.

As far as the schwa definition is concerned, it is a particular vowel: it can be produced or not, depending on several criteria like the communication situation, speaker's geographical origin etc. For example, a French speaker can say:

- "*J'ai pris le bus pour venir à la conférence*" [ʒepʀiləbyspuʀvəniʀalakɔ̃ferɑ̃s] (I took the bus to come to the conference)
- or "*J'ai pris l(e) bus pour v(e)nir à la conférence*" [ʒepʀilbyspuʀvniʀalakɔ̃ferɑ̃s].

Four main contexts can be distinguished. In this study, we only take into account the cases of variable elision in monosyllabic words, so when the monosyllabic word precedes a word starting with a consonant. So, we do not deal with contexts such "*l'ami*" [lami] (the friend) where the schwa elision is obligatory in the article "*le*".

Child-directed speech and acquisition of phonological variables

Compared to adult-directed speech (ADS), CDS appears to be more associative, repetitive and consistent (Hills, 2013). At the syntactic and lexical levels, we can note that CDS contains shorter and simpler utterances, and a more limited lexicon than ADS (for a review: Snow & Ferguson, 1977 ; Fletcher & MacWhinney, 1995). As for the phonological level, parents tend to speak slower to their children in a higher pitched-voice and with exaggerated intonation. Concerning phonological variables, Foulkes *et al.* (2005) note that parents tend to use standard variants more frequently than vernacular variants when they address their children.

Limited attention has been given to the two phonological variables we deal with in this study. The only study about schwa elision is the Phd work of Andreassen. She notes that there are less schwa elisions in CDS, mainly in nouns (Andreassen, 2007). We think that these phonological specificities show the speaker's wish to respect the standard and recurrent Consonant-Vowel structure of the syllable in French. The two phonological variables that we study are phenomena that cause problems of boundaries. Indeed, lexical boundaries and syllable boundaries may not correspond, depending on the production or omission of the phonological variables. For example when a French speaker says "*un ours*" (a bear), the liaison consonant is attached to the Word2 and forms a syllable containing the liaison consonant and the standard form of the Word2. Schwa elision also causes a problem. For example, there are two possible productions of the nominal phrase "*le chat*" (the cat). When the non-schwa variant of the article is used, it forms a syllable with the standard form of the noun "*l(e) chat*" [lʃa]. But when the standard variant is used, the article and the noun form two different syllables: [lə] and [ʃa].

To maximise speech segmentation, comprehension and lexical acquisition, parents tend to respect the Consonant-Vowel schema recurrent in French. In a previous study, we found that variable liaisons are produced more frequently in CDS than in ADS (Liégeois, Chabanal, & Chanier, 2011). Concerning the schwa in nouns, it is more often produced in CDS than in ADS (Andreassen, 2007).

The child, since an early age, is very sensitive to the regularities of syllable structures in the first language. So the lexical segmentation made by the child will depend on production variability. That's what Chevrot *et al.* (2009) demonstrated studying liaison acquisition. Within a usage-based framework of language acquisition (Barlow & Kemmer, 2000; Tomasello, 2003), they proposed a scenario of categorical liaison acquisition, divided in three steps (see Table 1). First, the child memorises a chunk containing the Word1, the liaison consonant and the Word2. At this stage of the development, the child does not produce any liaison errors. At step 2, errors appear. On the basis of syllable boundaries, the child memorises several exemplars of the same word. For example, the child can memorise the exemplar "*n-ours"* [nurs], on the basis of the production of "*un ours*" [œ̃nurs]. He also memorises the exemplars z-ours [zurs], t-ours [turs] etcetera. So, the child makes a mistake inserting the wrong exemplar of the Word2 after a particular Word1 (for example "*un z-ours*" [œ̃zurs]).

At step 3, the child has constructed a more abstract pattern of production. He has heard a lot of Word2 produced after a particular Word1, and he has learnt to produce the good exemplar of a Word2 after a particular Word1. At this stage, we do not observe any liaison errors, except cases of overgeneralisation of this pattern use (for example "*un nèbre*" [œ̃nɛbʀ] instead of "*un zèbre*" [œ̃zɛbʀ]).

Step	Input	Memorisation
Step 1	*un ours*	[un n – ours]
Step 2	*un ours*, *des ours*, *petit ours*	[n – ours] [z – ours] [t – ours] Example of error: un –z– ours
Step 3	*un ours*, *un âne* (a donkey), *un enfant* (a child)	[un – n+X]

Table 1: Scenario of categorical liaison acquisition (Chevrot et al., 2009)

Our main research question is as follows: Do CDS characteristics impact on the acquisition of the variable schwa? In this study, we focus only on schwa elision in monosyllabic words, which raise the boundary problems presented earlier. We deal with the closed set of the monosyllabic words *ce*, *de*, *je*, *le*, *me*, *ne*, *que*, *se* and *te*. To achieve this goal, we analysed parent-children interactions.

ALIPE project: data collection methodology and data structuration

In this study, we analyse the productions of three children and their parents, recorded during everyday situations (bath time, meals, playtime…) for one week. Parents had to record interactions for about one hour per day and make new recordings eight months later. This method has the advantage of being unobtrusive, as the researcher is not present during the recording. Moreover, this methodology requires few parental interventions, because they just had to start and stop the recorder. We present the corpus in the Table 2 below:

	Salomé	**Baptiste**	**Prune**
Age at stage 1 (S1)	28 months	36 months	40 months
Age at stage 2 (S2)	36 months	44 months	48 months
Coverage: length of recordings	14h (liaison) 10h (schwa)	10h (liaison + schwa)	14h (liaison) 11h (schwa)
Coverage: number of liaisons contexts	2347	1555	3363
Coverage: number of monosyllabic schwas contexts	6126	2667	6007

Table 2: corpus description

We recorded 2 small girls (Salomé and Prune) and one boy. Analyses about schwa elision are based on 31 hours of recording, about 10 hours per family. The results of schwa elision in monosyllabic words are based on six thousands contexts in the data of the girls, and about two thousands six hundreds for Baptiste's family. Our recordings have been transcribed and annotated the XML format, and transcription conventions especially designed for the ALIPE project. We have adopted an orthographic transcription, but used some codes to annotate liaison and schwa elision. We also annotated, for each parental utterance, if the adult talked to another adult or to the child.

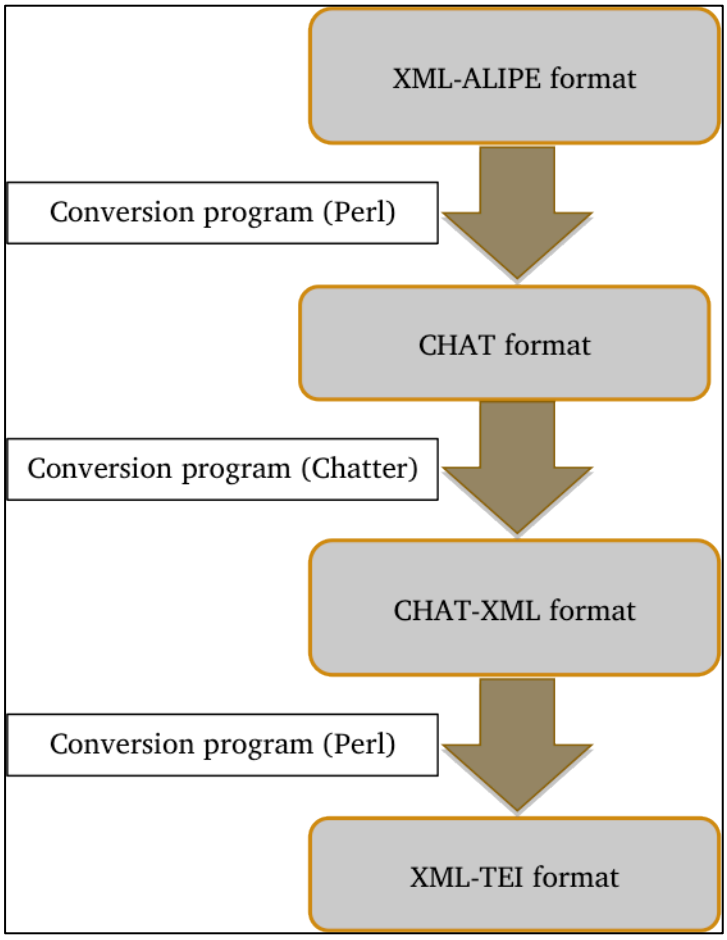

Figure 1: corpus structuration and transformation

By transformations using programs that we have designed and Chatter, a conversion program available on the website of the CHILDES database (MacWhinney & Snow, 1985; MacWhinney, 2000), we obtain 2 standard formats of our corpora: the CHAT format and the XML-TEI format (see Figure 1). The CHAT format, linked to the CLAN program, allows us to make automatic analyses like MLU. The use of these standards is important especially because it enables data exchange between researchers. Moreover, it allows other researchers not involved in the ALIPE project to contribute to the corpora's enrichment (by adding a new tier of annotation, for example), the replication of analyses and the verification of results.

All of our data are freely available as open-source data. We have designed a database to allow other users to download our transcriptions with the audio files (lrl-diffusion.univ-bpclermont.fr/). We also project to deposit our CHAT versions on the CHILDES database.

Acquisition of elision: what do we learn from CDS?

Based on the data presented above, we compared elision rates in parents' productions concerning all the contexts of monosyllabic words. In the Table 3, for each stage of recording, parents' productions are divided into two categories: ADS and CDS. For all parents, at S1, we can see that elision rates are significantly different in ADS and in CDS. At an early stage, parents tend to produce more often the schwa variant of the monosyllabic word when they address their children.

At S2, Salomé's parents elide in the same proportion when they address their child or another adult. The difference between ADS and CDS is no more significant. We can observe the same thing concerning Prune's parents. As for Baptiste's parents, the situation is different. The difference in elision rates in CDS and ADS at both stages remains significant. If we compare those rates with the ones in the children's productions (see Table 4), we can see that the elision rates of the two little girls increase between S1 and S2 while Baptiste elides in the same proportions at the two stages.

Parents	Stage of recording and age	Rate of elision in ADS	Rate of elision in CDS	χ^2	P
Salomé's parents	T1 : 2;4 ans	62,10%	37,00%	Chi2=95.0865	p<0,0001
	T2 : 3;0 ans	58,20%	56,10%	Chi2=0	p>0,05
Baptiste's parents	T1 : 3;0 ans	65,1%	31,7%	Chi2=75.9812	p<0,0001
	T2 : 3;7 ans	57,5%	35,9%	Chi2=40.7326	p<0,0001
Prune's parents	T1 : 3;4 ans	67,8%	31,6%	Chi2=95.0865	p<0,0001
	T2 : 4;0 ans	50,0%	51,2%	Chi2=0	p>0,05

Table 3: Rates of elision in ADS and CDS

Concerning schwa elision, we can say that parents tend to adapt their speech to the linguistic abilities of their child. At stage one, parents appear to maintain the schwa more frequently when they address their child. At stage two, however, the two girls (Salomé and Prune) produce elision rates that are comparable to the parents. The parents adjust their speech and elide similarly in CDS and ADS.

Child	Stage of recording and age	Rate of elision	χ^2	P
Salomé	T1 : 2;4 ans	10,81%	χ^2 = 216.0006	P<0,0001
	T2 : 3;0 ans	44,91%	Chi2=0	p>0,05
Baptiste	T1 : 3;0 ans	9,2%	χ^2 = 0.0527	P>0,05
	T2 : 3;7 ans	10,4%	Chi2=40.7326	p<0,0001
Prune	T1 : 3;4 ans	44,19%	χ^2 = 7.3637	P<0,01
	T2 : 4;0 ans	51,9%	Chi2=0	p>0,05

Table 4: elision rates in children's productions

The next analyses we present concern schwa elision in two monosyllabic words : "*le*" ("the") and "*je*" ("I"). The monosyllabic "*le*" is a special case as it can be used like a pronoun (preceding a verb) or like an article (preceding a noun phrase). In the continuation of this study, we only focus on "*le*" article and "*je*" to determine whether the phonological variation in these contexts is modulated in CDS and whether it has an impact on the acquisition of elision.

As far as the monosyllabic "*le*" is concerned, we analyse all the contexts where it is used like an article. We identified three different construction categories depending on the way the schwa is elided or maintained. Let's take the example of "*le ballon*" (the ball), which is a construction. This particular construction can be categorised as:

- A retention construction if the speaker always uses the schwa variant of the article within this construction: [l ə b a l õ].
- An elision construction if the speaker always uses the non-schwa variant of the article preceding "*ballon*": [l b a l õ].
- A variable construction when variability is observed between the schwa and the non-schwa variant in the speaker's productions.

Table 5 below presents, for each speaker, the number of constructions on which our analyses are based.

Child	**Number of constructions (S1 + S2)**	**Parents**	**Number of constructions (S1 + S2)**
Salomé	117	Salomé's parents	180
Baptiste	81	Baptiste's parents	126
Prune	117	Prune's parents	96

Table 5: number of constructions analysed

At each stage of recording, we calculated the proportion of retention, elision, and variable contexts both in CDS and in the children's productions. In Salomé's productions (see Figure 2), we note that retention of the schwa is almost systematic. At stage two, we observe an increase in this proportion, and this observation is also valid for the parents' productions. Concerning elision constructions, they are marginal in the child's productions at S1. They represent only about 5% of the total occurrences. As to variable contexts, there are inexistent at stage 1. The CDS tend to follow child's productions tendencies. Between S1 and S2, parents produce more often elision constructions than retention and variable constructions.

Concerning Baptiste (see Figure 3), we note that the proportions in his productions do not vary a lot between the two stages of recording. Like Salomé, he does not produce any variable contexts at S1, and very few elision contexts (only 2%). At S2, retention constructions are still dominant and they represent more than 82%. Between the two stages of recording, the characteristics in CDS do not present major changes.

For its part, Prune produces less retention constructions than the two other children (see Figure 4). It seems logical because she is the oldest child. Concerning elision and variable constructions, we can note an increase, both in the child's and in the parents' productions.

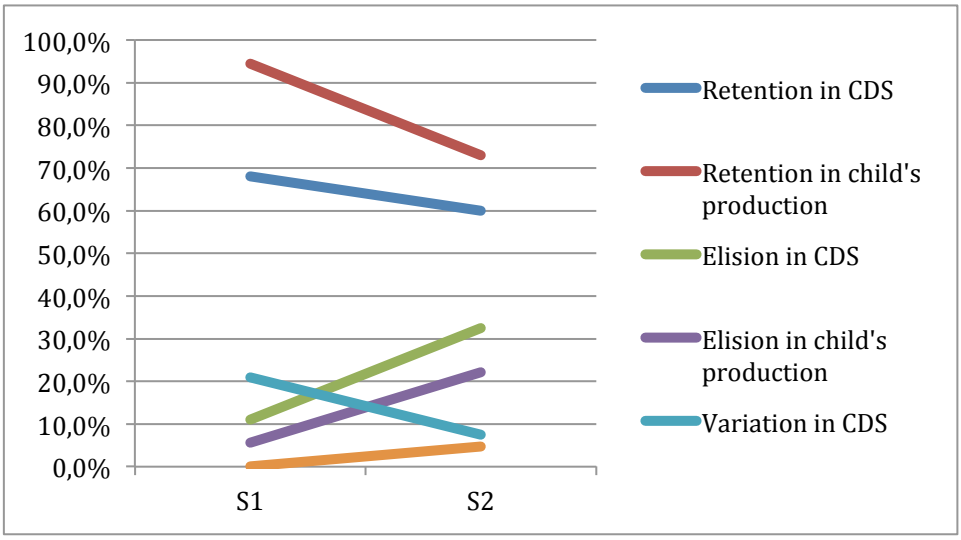

Figure 2: proportions of construction categories in the productions of Salomé and her parents

To summarise, we can say that, at an early stage of acquisition, elision is marginal in the context le + a noun phrase. At this early stage, children mainly use the standard variant of "*le*" before a noun or an adjective. Since variable contexts are inexistent in the children's productions, we postulate that the article is not memorised with a variable schwa. But, since we observe a few constructions where the article "*le*" is produced with the non-schwa variant, we suggest that the non-schwa variant of the article is memorised as a chunk with the noun, like liaison variants, to form a complex onset of the word, with two consonants.

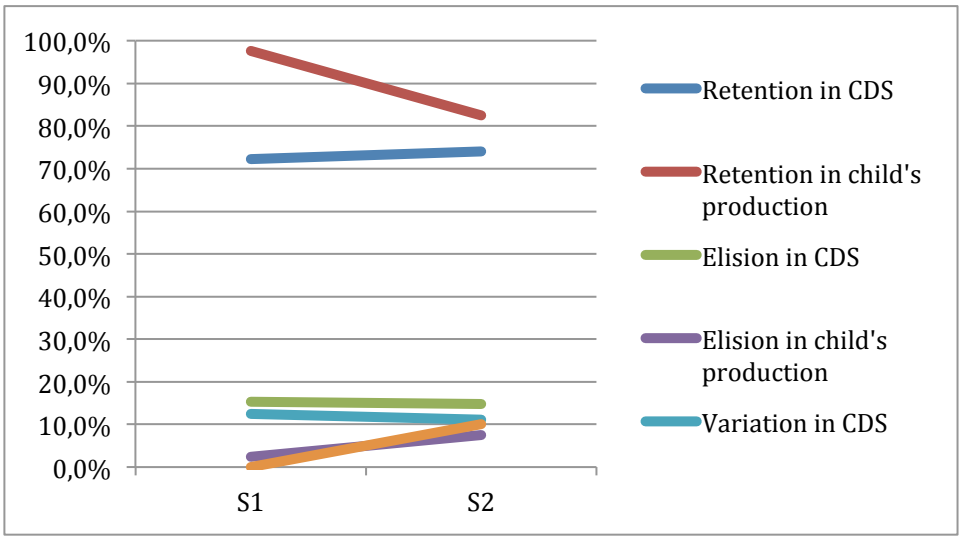

Figure 3: proportions of construction categories in the productions of Baptiste and his parents

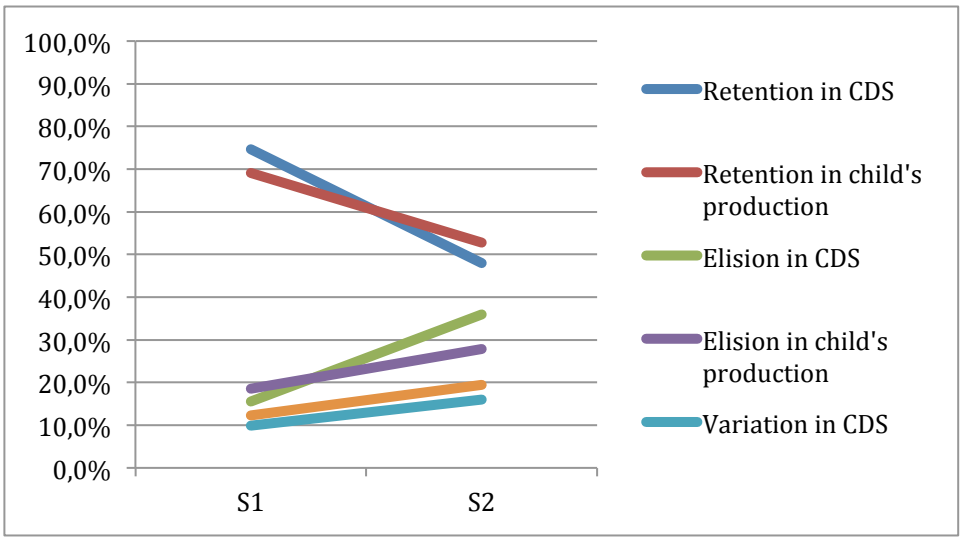

Figure 4: proportions of construction categories in the productions of Prune and her parents

The second monosyllabic word we would like to focus on is "*je*". Since constructions with "*je*" are more frequent than with the monosyllabic word "*le*", we can observe elision rates in specific constructions, without construct categories. For each family, we have chosen the specific constructions that appear more than twenty times in the productions of the children at S2 and in the CDS at S1. Unfortunately, we cannot study the case of Baptiste because he doesn't talk much and he uses very infrequently the pronoun "je". He is a special case because he uses almost exclusively the third person singular or the complement pronoun "*moi*" ("me") when he speaks about himself. This tendency of the child to use the third person when she talks about herself has been studied by Morgenstern (2010).

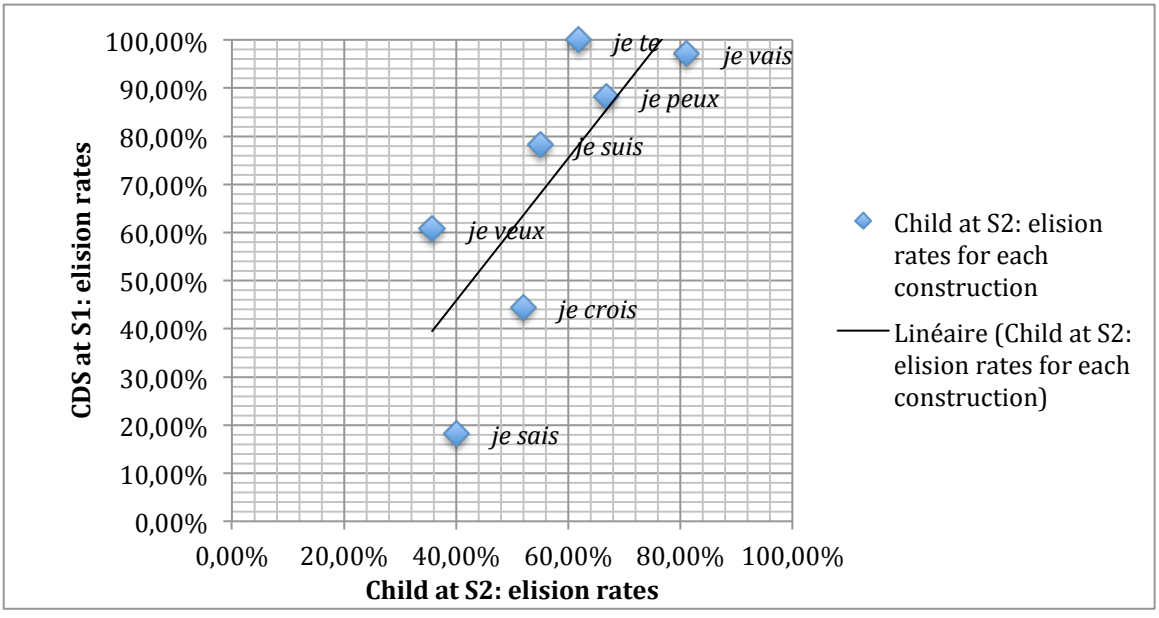

Figure 5: elision rates for each construcion in Salomé's productions at S2 and CDS at S1

We first observe Salomé's productions (see Figure 5). We can see that elision rates at S1 in CDS are correlated with elision rates in Salomé's productions at S2, using the Spearman rank correlation test (Rho = 0.786, p<0,05). We can say that constructions that are more frequently concerned by elision in CDS at S1 are constructions that are more frequently concerned by elision in Salomé's productions at S2. The same observations are true concerning Prune's productions (see Figure 6). Constructions that are more frequently concerned by elision in CDS at S1 are constructions that are more frequently concerned by elision in Prune's productions at S2 (Rho = 1, p<0,05).

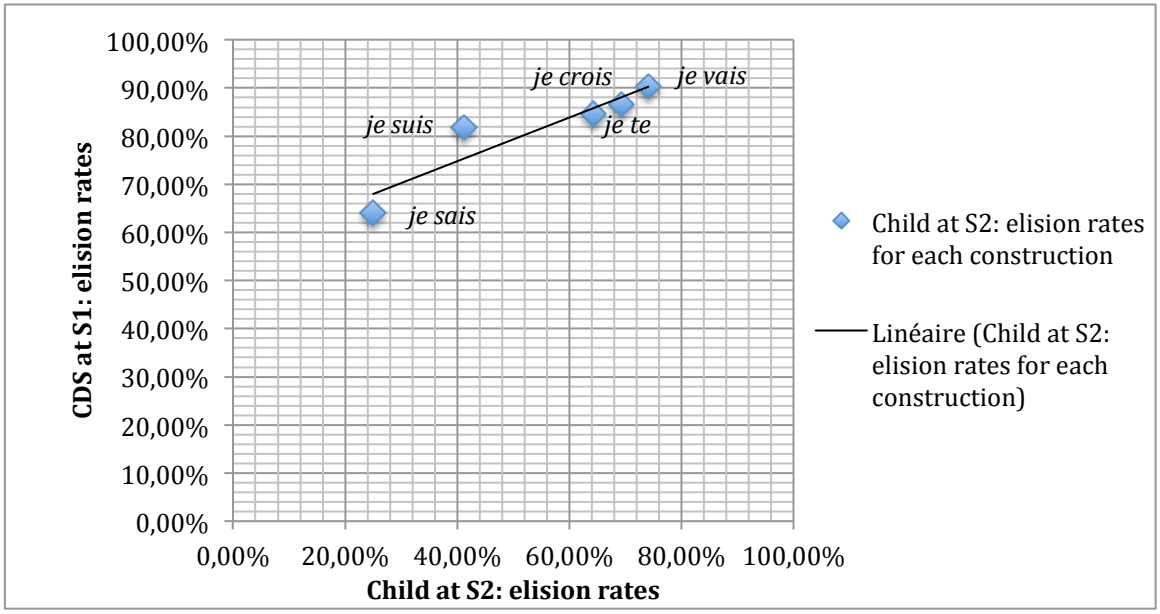

Figure 6: elision rates for each construcion in Salomé's productions at S2 and CDS at S1

Discussion

In comparison to ADS, we can conclude that CDS is adjusted concerning phonological variation. Variation not only concerns syntax, lexicon and prosody but phonological variables too. At an early acquisition stage, rates of elision are significantly different between CDS and ADS. We tested other variables like speech rates and the phoneme following the monosyllabic words and they don't predict this difference.

For the particular constructions [le + NOUN GROUP], we noted very low elision rates in CDS. We think that parents tend to clearly separate the article and the noun or the adjective because of their functional predominance (Andreassen, 2007), respecting the recurrent pattern of the syllable in French. What's more in these contexts, children don't vary their productions at an early stage. They predominantly use the schwa variant of the article "*le*", and they don't vary between the schwa and the non-schwa variant.

These findings allow us to propose a scenario of acquisition for this context. The absence of variability in children's productions suggests that the article is not memorized with a variable schwa at an early stage. We propose that the child memorises two types of constructions:

- An exemplar containing the non-schwa variant of the article with the substantive, e.g. "*l-chat*" [l ʃ a] ("the cat").

- A more general pattern whose permit her to insert a noun or an adjective beginning with a consonant after "*le*": [le + X]

Later, the child is exposed to a more variable input and is able to construct and memorise a more abstract pattern allowing him to vary her production, in the same context, between the schwa variant and the non-schwa variant. We can also postulate that two different patterns stay in competition. So, both the schwa and the non-schwa variants of the article would be stored as lexical entries, like Bürki *et al*. (2011; 2010) propose for nouns.

The observation of three children, and using only two stages of observation are not sufficient. These findings need to be checked by studying other corpora or through experimental tests. Other constructions concerned by schwa elision also have to be studied.

Bibliography

Andreassen, H. N. (2007). La distinction /ø/ - /ə/ dans l'acquisition : input et output chez des enfants suisses. *Actes des JEL 2007 (Journées d'Etudes Linguistiques)* (pp. 77–82).

Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage Based Models of Language*. Center for the Study of Language and Information.

Bürki, A., Alario, F. X., & Frauenfelder, U. H. (2011). Lexical representation of phonological variants: Evidence from pseudohomophone effects in different regiolects. *Journal of Memory and Language*, *64*(4), pp. 424–442. doi:10.1016/j.jml.2011.01.002

Bürki, A., Fougeron, C., Gendrot, C., & Frauenfelder, U. H. (2010). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics*, *39*(3), pp. 279–288. doi:10.1016/j.wocn.2010.07.003

Chevrot, J. P., Dugua, C., & Fayol, M. (2009). Liaison acquisition, word segmentation and construction in French: a usage based account. *Journal of child language*, *36*(3), pp. 557–596.

Fletcher, P., & MacWhinney, B. (Eds.). (1995). *The Handbook of Child Language*. Hoboken: Wiley-Blackwell.

Foulkes, P., Docherty, G. J., & Watt, D. (2005). Phonological Variation in Child-Directed Speech. *Language*, *81*(1), pp. 177–206. doi:10.1353/lan.2005.0018

Hills, T. (2013). The company that words keep: comparing the statistical structure of child- versus adult-directed language. *Journal of child language*, *40*(3), pp. 586–604. doi:10.1017/S0305000912000165

Liégeois, L., Chabanal, D., & Chanier, T. (2011). La liaison en discours adressé à l'enfant, spécificités et impacts sur l'acquisition. *Colloque du Réseau Français de Phonologie*. 1-3 juillet 2011, Tours.

MacWhinney, B. (2000). *The CHILDES project: Tools for Analyzing Talk*. (L. E. Associates, Ed.) (3rd ed.). Mahwah, NJ.

MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of Child Language*, *12*(2), pp. 271–295.

Morgenstern, A. (2010). JE comme AUTRE Mise en mots de l'altérité et de l'identité dans le langage de l'enfant. In L. Dufaye & L. Gournay (Eds.), *L'altérité dans les théories de l'énonciation* (pp. 115–135).

Snow, C. E., & Ferguson, C. A. (Eds.). (1977). *Talking to Children: Language Input and Acquisition*. Cambridge University Press.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.